

Textos Explicativos

Probabilidade e Estatística

Material de apoio para a Lista de Exercícios — Probabilidade (espaço amostral, eventos, união, independência, condicional);
média, mediana, moda, média ponderada, variância e desvio padrão

1. Probabilidade Clássica e Evento Complementar

Questões relacionadas: 1 e 2

Probabilidade é a medida da chance de um evento acontecer. O conjunto de todos os resultados possíveis de um experimento aleatório é o espaço amostral (Ω), e um evento (A) é o subconjunto dos resultados que nos interessam. Quando todos os resultados são igualmente prováveis, a probabilidade é a razão entre os casos favoráveis e os casos possíveis, e vale sempre entre 0 (impossível) e 1 (certo):

$$P(A) = n(A) / n(\Omega) \quad P(\text{não } A) = 1 - P(A)$$

A segunda fórmula é a do evento complementar: a probabilidade de um evento não ocorrer é o que falta para chegar a 1. Ela é muito útil quando é mais fácil calcular "o que não queremos".

Exemplo 1 — dado: ao lançar um dado comum, $\Omega = \{1, 2, 3, 4, 5, 6\}$, logo $n(\Omega) = 6$. Número par: $\{2, 4, 6\}$, então $P = 3/6 = 1/2 = 50\%$. Número maior que 4: $\{5, 6\}$, então $P = 2/6 = 1/3 \approx 33,3\%$.

Exemplo 2 — urna: uma urna tem 5 bolas vermelhas, 3 azuis e 2 verdes (total 10). $P(\text{azul}) = 3/10 = 30\%$. Para "não vermelha": há $3 + 2 = 5$ bolas favoráveis, então $P = 5/10 = 1/2$. Pelo complementar, o resultado é o mesmo: $1 - 5/10 = 1/2$.

2. Probabilidade da União de Eventos ("A ou B")

Questão relacionada: 3

Para calcular a probabilidade de ocorrer o evento A ou o evento B, somam-se as duas probabilidades. Porém, se os eventos têm resultados em comum (a interseção $A \cap B$), esses resultados foram contados duas vezes, e é preciso descontá-los uma vez:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Quando os eventos não têm nada em comum (são mutuamente exclusivos), $P(A \cap B) = 0$ e basta somar.

Exemplo prático: de um baralho de 52 cartas, qual a probabilidade de sair um rei ou uma carta de copas? $P(\text{rei}) = 4/52$ e $P(\text{copas}) = 13/52$. O rei de copas pertence aos dois eventos, então $P(\text{rei e copas}) = 1/52$. Logo $P(\text{rei ou copas}) = 4/52 + 13/52 - 1/52 = 16/52 = 4/13 \approx 30,8\%$.

3. Interseção de Eventos ("A e B"): Independentes e Condicional

Questões relacionadas: 4 e 5

Para calcular a probabilidade de dois eventos ocorrerem em sequência (A e B), multiplicam-se as probabilidades. Há dois casos. Se os eventos são independentes — o resultado de um não altera a probabilidade do outro, como em lançamentos de moeda —, basta multiplicar as probabilidades de cada um. Se são dependentes — como em retiradas sem reposição —, a probabilidade do segundo evento deve ser calculada sabendo que o primeiro já ocorreu, a chamada probabilidade condicional $P(B | A)$:

$$\text{Independentes: } P(A \cap B) = P(A) \cdot P(B) \quad \text{Dependentes: } P(A \cap B) = P(A) \cdot P(B | A)$$

Exemplo 1 — moeda (independentes): uma moeda honesta é lançada 3 vezes. Em cada lançamento, $P(\text{cara}) = 1/2$. $P(3 \text{ caras}) = 1/2 \times 1/2 \times 1/2 = 1/8 = 12,5\%$. Para "pelo menos uma coroa", usa-se o complementar: é o contrário de "nenhuma coroa" (3 caras), então $P = 1 - 1/8 = 7/8 = 87,5\%$.

Exemplo 2 — urna sem reposição (condicional): uma urna tem 4 bolas vermelhas e 6 azuis (total 10). Tirando duas bolas, uma após a outra, sem reposição, qual a probabilidade de ambas serem vermelhas? A 1ª é vermelha com $P =$

4/10. Como a bola não volta, restam 3 vermelhas em 9 bolas: $P(2^{\text{a}} \text{ vermelha} \mid 1^{\text{a}} \text{ vermelha}) = 3/9$. Logo $P = 4/10 \times 3/9 = 12/90 = 2/15 \approx 13,3\%$. Se houvesse reposição, os eventos seriam independentes e a conta seria $4/10 \times 4/10 = 16\%$.

4. Média, Mediana e Moda

Questões relacionadas: 6 e 7

Média, mediana e moda são as medidas de tendência central: cada uma resume um conjunto de dados em um único valor "típico". A média aritmética é a soma dos valores dividida pela quantidade de valores. A mediana é o valor central dos dados em ordem crescente (se a quantidade de dados é par, é a média dos dois valores centrais). A moda é o valor que mais se repete — o de maior frequência.

$$\text{Média: } \bar{x} = (x_1 + x_2 + \dots + x_n) / n$$

Exemplo 1 — média: as notas de um aluno foram 6,0; 7,5; 8,0; 5,5 e 9,0. A soma é 36,0, então $\bar{x} = 36,0/5 = 7,2$.

Exemplo 2 — média, mediana e moda: no conjunto 3, 5, 7, 5, 9, 5, 8, 2, a soma é 44 e $n = 8$, então $\bar{x} = 44/8 = 5,5$. Em ordem crescente: 2, 3, 5, 5, 5, 7, 8, 9. Como n é par, a mediana é a média do 4º e do 5º valores: $(5 + 5)/2 = 5$. O valor 5 aparece 3 vezes, mais do que qualquer outro, então a moda é 5.

5. Média Ponderada e Média com Tabela de Frequências

Questões relacionadas: 8 e 9

Na média ponderada, cada valor tem um peso, que indica sua importância no resultado. Multiplica-se cada valor pelo seu peso, soma-se tudo e divide-se pela soma dos pesos:

$$\bar{x}_p = (p_1 \cdot x_1 + p_2 \cdot x_2 + \dots + p_n \cdot x_n) / (p_1 + p_2 + \dots + p_n)$$

Exemplo 1 — nota final: a Prova 1 tem peso 2, a Prova 2 tem peso 3 e o Trabalho tem peso 1. Um aluno tirou 6,0, 8,0 e 9,0. A soma dos pesos é $2 + 3 + 1 = 6$, e a soma ponderada é $2 \times 6,0 + 3 \times 8,0 + 1 \times 9,0 = 12 + 24 + 9 = 45$. A nota final é $45/6 = 7,5$.

A tabela de frequências segue a mesma ideia: a frequência (f_i) de cada valor funciona como seu peso, e a soma das frequências é o total de dados (n). A moda é o valor de maior frequência, e a frequência relativa mostra a fração do total que cada valor representa:

$$\bar{x} = \sum(f_i \cdot x_i) / n \quad f_r = f_i / n \times 100\%$$

Exemplo 2 — idades de uma turma: 15 anos — 4 alunos; 16 anos — 10 alunos; 17 anos — 6 alunos ($n = 20$). Média: $\bar{x} = (15 \times 4 + 16 \times 10 + 17 \times 6) / 20 = (60 + 160 + 102) / 20 = 322 / 20 = 16,1$ anos. Moda: 16 anos (maior frequência, 10 alunos). Frequência relativa dos alunos de 16 anos: $f_r = 10/20 = 0,5 = 50\%$.

6. Variância e Desvio Padrão

Questão relacionada: 10

A média resume o centro dos dados, mas não diz o quanto eles são espalhados. Duas turmas podem ter a mesma nota média e, mesmo assim, uma ter notas todas parecidas e a outra ter notas muito diferentes entre si. A variância e o desvio padrão medem essa dispersão em torno da média: quanto menores, mais concentrados estão os dados.

A variância é a média dos quadrados dos desvios (a diferença de cada valor para a média). O desvio padrão é a raiz quadrada da variância, o que o traz de volta à mesma unidade dos dados:

$$\sigma^2 = \sum(x_i - \bar{x})^2 / n \quad \sigma = \sqrt{\sigma^2}$$

Exemplo prático: para os dados 4, 6, 8, 10 e 12: a média é $40/5 = 8$. Os desvios são $-4, -2, 0, 2$ e 4 , e seus quadrados são 16, 4, 0, 4 e 16, com soma 40. A variância é $\sigma^2 = 40/5 = 8$, e o desvio padrão é $\sigma = \sqrt{8} \approx 2,83$.